Letter to the Editor

Patterns of Diversity Among SINE Elements Isolated from Three Y-Chromosome Genes in Carnivores

Jill Pecon Slattery, William J. Murphy, and *Stephen J. O'Brien*

Laboratory of Genomic Diversity, Frederick Cancer and Research Development Center, National Cancer Institute, Frederick, Maryland

Short interspersed elements (SINEs) are forms of ''selfish'' DNA scattered throughout eukaryotic genomes (Jelinek and Schmid 1982; Daniels and Deininger 1985). SINEs exist in high copy numbers and, combined with long interspersed nuclear element (LINEs) and retroelements resembling retroviruses, may constitute 36% of the total genome (Smit et al. 1996). A consensus of research of the human *Alu* SINE family and novel SINEs from other vertebrate and invertebrate taxa (see Okada 1991*a*, 1991*b*; Ohshima and Okada 1994; Shimamura et al. 1997) defines evolutionary lineages, or subfamilies, which originated from small RNA molecules such as 7SL RNA (*Alu*) or tRNAs. Proposed mechanisms governing the proliferation of SINE subfamilies generally invoke transcription in high copy number of a source, or master copy, SINE, followed by reverse transcription into DNA and reinsertion at a new site (Deininger et al. 1992; Brookfield 1994; Smit 1996; Kidwell and Lisch 1998), perhaps selected by specific target sequence (Jurka and Klonowski 1996; Tatout, Lavie, and Deragon 1998). The possible functions of SINE elements within the genome remain unclear, yet studies suggest a role in gene and oncogene expression (Britten 1996, 1997), recombination between nonhomologous regions, and mutagenesis (Amariglio and Rechavi 1993; Michel et al. 1997). Although research of primates describes a high incidence of SINE elements within the Y chromosome (Smith et al. 1987; Baird and Royle 1997), little is known of the evolution of these repetitive sequences in the nonrecombining region (NRY). Here, we provide a unique phylogenetic depiction of SINEs isolated from intronic regions of three genes located in the NRY. Using representative taxa from the Order Carnivore, we demonstrate the following: (1) major SINE lineages exist specific to each carnivore family; (2) within carnivore families, SINE retroposons provide cladistic markers of speciation, but also (3) exhibit stochastic intraspecies transposition; and (4) homoplasy exists with identical insertions in two species from distantly related lineages.

Intron segments from *Smcy*, *Zfy*, and *Ube1y* were sequenced from 20 species of Felidae and 8 taxa representing 6 additional carnivore families (GenBank accession numbers AF111609–AF221621). The order Carnivora consists of two major superfamilies: catlike carnivores (Feloidea), including the Felidae (cats), Viverridae (viverrids), Herpestidae (mongooses), and Hyaenidae (hyenas) families, and doglike carnivores (Arctoidea), composed of the families Canidae (dogs), Ursidae (bears), Mustelidae (skunks, weasels), Procyonidae (raccoons), and pinnipeds (seals, walruses). Twenty of the 37 species within the cat family, representing distinct evolutionary lineages (fig. 1*A*) established from a consensus of multiple genetic markers (Collier and O'Brien 1985; Modi and O'Brien 1988; Pecon Slattery et al. 1994; Janczewski et al. 1995; Johnson et al. 1996; Masuda et al. 1996; Johnson and O'Brien 1997; Pecon Slattery and O'Brien 1998), were sequenced. Felid SINE retroposons were sporadic in the three genes and limited to the seven species within the domestic cat lineage and two other distantly related species of *Lynx rufus* and *Otocolobus manul* (fig. 1*A*). The remaining SINEs occurred in only three of the six families examined and all from the Arctoidea: Canidae (*Canis familiaris,* domestic dog), Procyonidae (*Ailurus fulgens,* red panda), Ursidae (*Ailuropoda melanoleuca,* giant panda; *Tremarctos ornata,* spectacled bear; and *Ursus arctos,* brown bear).

Sequence similarity with published mammalian SINE inserts supported an evolutionary model in which SINEs were derived, in part, from tRNA. The newly identified carnivore SINEs possessed a 5′ region containing the split block promoter for RNA polymerase III common to tRNA genes, followed by a unique sequence and a dinucleotide repeat of variable length, and terminated by a 3′ polyA tail (fig. 1*B*). As with published B2-like SINEs from the mink X chromosome (Mvi B2), dog (Cfa B2) and harbor seal (Pvi B2) sequence similarity (fig. 1*B*) indicated that the newly described carnivore SINEs likely originated from tRNA-*lys* or tRNA-*arg* (Lavrentieva et al. [1991] and Coltman and Wright [1994], respectively).

In addition, our findings demonstrate the presence of taxon-specific SINE subfamilies within carnivores. Unique to the cat, bear, and raccoon families, we propose these newly described SINEs as subfamilies. In the domestic dog (Canidae), a higher level of genetic distance between the *Zfy* SINE and that of the published canid B2-like SINE is represented by the placement in different lineages within the SINE phylogeny (fig. 2). Furthermore, a lack of sequence similarity with published SINE flanking regions for the domestic dog (Das et al. 1998) indicated that the *Zfy* SINE diverged from a master copy as yet not defined. Thus, the *Zfy* SINE may be either the remnant of an extinct lineage or a more recent version of the canid B2-like SINE. Moreover, the low level of similarity between autosomal and Y chromosome SINE inserts observed here in the do-

Key words: retroposon, SINE, carnivore, Y chromosome, evolution.

Address for correspondence and reprints: Jill Pecon Slattery, Laboratory of Genomic Diversity, Frederick Cancer and Research Development Center, National Cancer Institute, Frederick, Maryland 21702. E-mail: slattery@mail.ncifcrf.gov.

Mol. Biol. Evol. 17(5):825–829. 2000

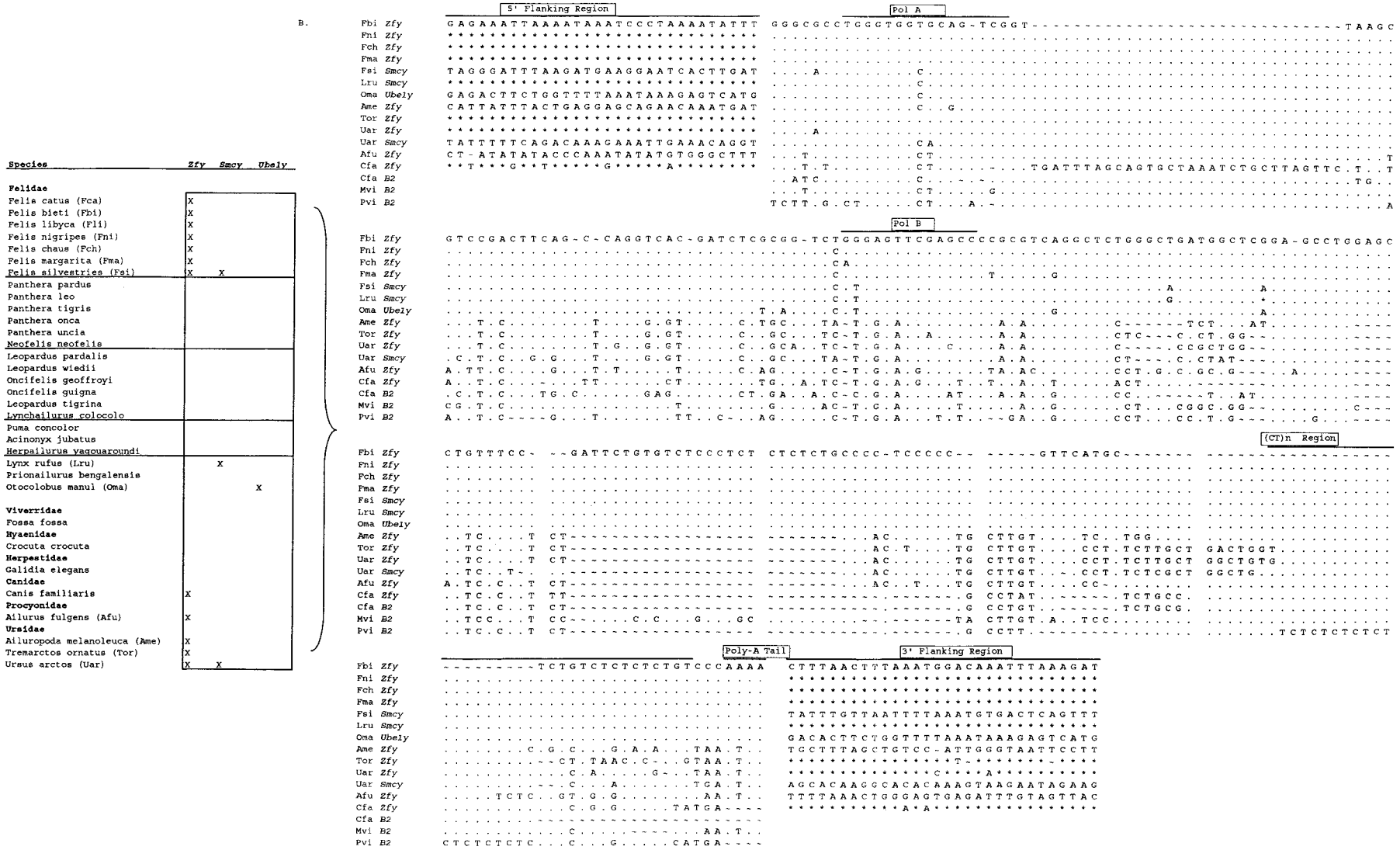

FIG. 1.—*A*, Species of carnivores sequenced for *Zfy*, *Smcy*, and *Ube1y* introns. Listed are family, species, lineage (Felidae only), and three-letter species code. Presence of the SINE is indicated by ''X''. *B*, Alignment of SINE sequences and flanking regions in carnivores. SINE sequences from *Zfy*, *Ube1y*, and *Smcy* introns were obtained via PCR of genomic DNA from a single male of each species (except for the domestic cat lineage, from which two individuals of *Fca, Fni*, *Fsi*, and *Fma* were examined) by methods described elsewhere (Pecon Slattery and O'Brien 1998; unpublished data). Sequences of B2-like elements were obtained from GenBank with accession numbers X52381 (Mvi), X57357 (Cfa), and Z33499 (Pvi). Identical SINEs with Fbi were isolated from Fca, Fli, and Fsi. Alignment was constructed via CLUSTAL X (Thompson et al. 1997) and verified visually. **Poly A tail** denotes starting position but is not shown in its entirety. For the flanking regions, an asterisk denotes an identical site with the closest sequence above. For the SINE alignment, ~ denotes a gap and . indicates an identical site.

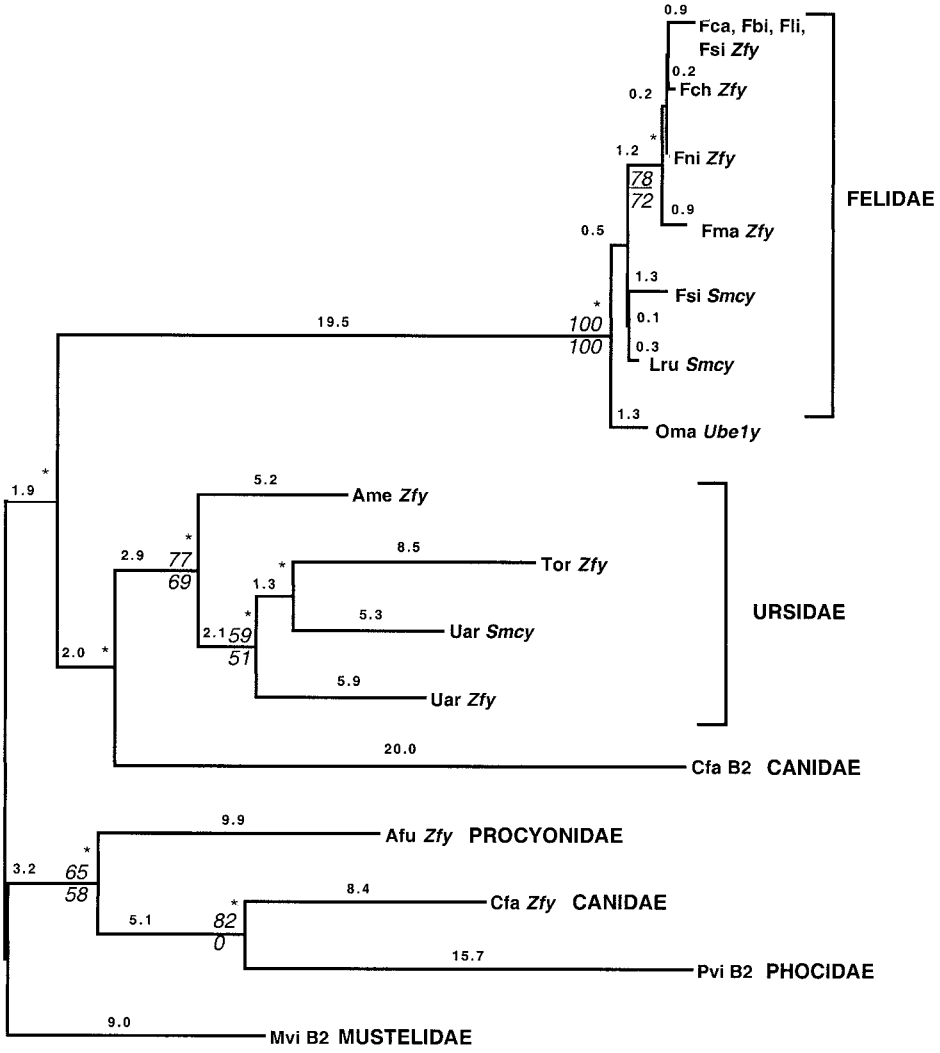

FIG. 2.—Phylogenetic reconstruction of SINE sequences from carnivore species. Shown is a tree constructed by minimum evolution estimated by the neighbor-joining (NJ) algorithm using the Tajima-Nei (Tajima and Nei 1984) model of substitution as implemented by PAUP, version 4.0b2 (used by permission from D. Swofford). A nearly identical topology (length = 427 steps; consistency index = 0.7869) is obtained from a 50% majority-rule consensus of 44 equivalent trees derived from maximum-parsimony (MP) analyses with PAUP. A nearly identical maximum-likelihood (ML) tree (−ln likelihood = 1,184.28; 524 trees examined) is obtained using PHYLIP, version 3.5 (Felsenstein 1993). Both MP and ML trees differ from the NJ tree shown here by slight rearrangements of internal branches within the Ursidae lineage. Branch lengths are percentages of sequence divergence (genetic distance × 100) computed by minimum evolution estimated by NJ with the Tajima-Nei model of substitution. Nodes supported by the ML analyses are indicated by asterisks. Bootstrap proportions (%) greater than 50% are based on 100 iterations with NJ above and MP values below.

mestic dog is also demonstrated in human *Alu* SINEs and may be a consequence of a different mechanism or rate of proliferation unique to the Y chromosome (Smith et al. 1987).

The pattern of SINE diversification within the well-characterized species phylogeny of Felidae provides new insights into the evolution of these repetitive sequences. First, the monophyletic grouping of SINEs from two different genes indicates a common ancestral copy unique to the cat family. Second, the SINE flanking regions (fig. 1*B*) are specific to each gene intron. In the case of *Zfy*, all seven species within the domestic cat lineage have the SINE element in the exact same location within the intron. Thus, the presence of the Zfy SINE is a cladistic synapomorphic trait marking an insertion event in a common ancestor to the extant species within the domestic cat lineage. Moreover, the SINE phylogeny accurately derives the expected evolutionary relationships by uniting *F. bieti* (Fbi), *F. catus* (Fca), *F. libyca* (Fli), and *F. silvestris* (Fsi) apart from *F. nigripes*, (Fni), *F. margarita* (Fma), and *F. chaus* (Fch). Therefore, this topology depicts SINE insertion subsequently followed by the gradual accumulation of mutations in accordance with known patterns of speciation.

In contrast, sporadic insertions unrelated to species divergence, as well as clear evidence of homoplasy, are represented by *Smcy* in Felidae. Only one species from the domestic cat lineage, *F. silvestris*, possessed a SINE within *Smcy*, an indication that this event was unique and occurred after the *Zfy* insertion. The presence of this SINE within the same location with identical flanking sequences (fig. 1*B*) of *Smcy* in *L. rufus*, a species only

distantly related to the domestic cat lineage, is likely an example of insertion dictated by the target sequences within the flanks. This provides strong evidence that SINE insertion at identical sites within different species can occur independently of phylogeny, and it counters the hypothesis that SINEs are exempt from parallel or convergent evolution (Nikaido, Rooney, and Okada 1999; for commentary, see Hillis 1999). Finally, a species-specific integration, found solely in *O. manul* for *Ube1y*, probably represents a recent insertion.

Similarly, the evolution of SINEs within *Zfy* and *Smcy* in Ursidae corroborates some of the patterns of diversification observed in Felidae. The presence of a SINE, in the same location within the *Zfy* intron for three species examined, denotes an ancestral synapomorphic insertion early within the bear phylogeny. Conversely, the presence of a *Smcy* SINE only in *U. arctos* suggests a recent, autoapomorphic insertion.

Across the broad taxonomic divisions represented by the carnivore families of Felidae, Ursidae, Procyonidae, Canidae, Mustelidae, and Phocidae, few sites are conserved in the SINE retroposon (fig. 1*B*). Of these, the most notable are the split block promoter sites of Pol A and Pol B, suggestive of functional constraints linked with binding RNA polymerase III. Within Felidae, little genetic diversity is detected, and genetic distance estimates vary from 0.5%–3.8% (not shown), independent of which gene intron harbored the SINE insert. In comparison, the four ursid SINEs from two genes are less conserved, with estimates of 12.9%–16.0% (not shown). Provided the ancestral source SINE occurred early within the evolution of each family, the differences between Ursidae and Felidae may be a consequence of an overall deeper divergence time of modern ursids (30 MYA) compared with modern Felidae (12–15 MYA) (Wayne, Van Valkenburgh, and O'Brien 1991).

The role of repetitive DNA on the Y chromosome is not clear. All three genes screened herein occur outside the pseudoautosomal region (Murphy et al. 1999) and reside in the sex-specific NRY. The fact that the X homolog of *Zfy* lacks the SINE element (Pecon Slattery and O'Brien 1998) in all domestic cat lineage species corroborates the prediction that repetitive DNA tends to accumulate in nonrecombining regions of the genome (Charlesworth 1991). The lack of conventional recombination in the NRY may provide the basis for a possible role of SINEs. As much of the eutherian Y chromosome undergoes extensive intrachromosomal shuffling (Graves 1995, 1998), it is possible that SINE retroposons may facilitate this process through nonhomologous recombination, a process modeled in the LINE-LINE recombination events which occurred on the Y chromosome during human evolution 4–5 MYA (Schwartz et al. 1998). Other possible functions for this ubiquitous form of repetitive DNA may be revealed by ongoing research into the evolution of coding regions of Y chromosome genes in Felidae.

LITERATURE CITED

AMARIGLIO, N., and G. RECHAVI. 1993. Insertional mutagenesis by transposable elements in the mammalian genome. Environ. Mol. Mutagen. **21**:212–218.

BAIRD, D. M., and N. J. ROYLE. 1997. Sequences from higher primates orthologous to the human Xp/Yp telomere junction reveal gross rearrangements and high levels of divergence. Hum. Mol. Genet. **6**:2291–2299.

BRITTEN, R. 1996. DNA sequence insertion and evolutionary variation in gene regulation. Proc. Natl. Acad. Sci. USA **93**: 9374–9477.

———. 1997. Mobile elements inserted in the distant past have taken on important functions. Gene **205**:177–182.

BROOKFIELD, J. F. 1994. The human Alu SINE sequences—is there a role for their selection in their evolution? Bioessays **16**:793–795.

CHARLESWORTH, B. 1991. The evolution of sex chromosomes. Science **251**:1030–1033.

COLLIER, G. E., and S. J. O'BRIEN. 1985. A molecular phylogeny of the Felidae: immunological distance. Evolution **39**:437–487.

COLTMAN, D. W., and J. M. WRIGHT. 1994. Can SINEs: a family of tRNA-derived retroposons specific to the superfamily Arctoidea. Nucleic Acids Res. **22**:2726–2730.

DANIELS, G. R., and P. L. DEININGER. 1985. Repeat sequence families derived from mammalian tRNA genes. Nature **317**: 819–822.

DAS, M., L. L. CHU, M. GHAHREMANI, T. ABRAMS-OGG, M. S. ROY, D. HOUSEMAN, and J. PELLETIER. 1997. Characterization of an abundant short interspersed nuclear element (SINE) present in *Canis familiaris.* Mamm. Genome **9**:64–69.

DEININGER, P. L., M. A. BATZER, C. A. HUTCHINSON III, and M. H. EDGELL. 1992. Master genes in mammalian repetitive DNA amplification. Trends Genet. **8**:307–311.

FELSENSTEIN, J. 1993. PHYLIP (phylogenetic inference package). Version 3.5. Distributed by the author, Department of Genetics, University of Washington, Seattle.

GRAVES, J. A. M. 1995. The origin and function of the mammalian Y chromosome and Y-borne genes—an evolving understanding. Bioassays **17**:11–20.

———. 1998. Evolution of the mammalian Y chromosome and sex-determining genes. J. Exp. Zool. **281**:472–481.

HILLIS, D. M. 1999. SINEs of the perfect character. Proc. Natl. Acad. Sci. USA **96**:9979–9981.

JANCZEWSKI, D., W. S. MODI, J. S. STEPHENS, and S. J. O'BRIEN. 1995. Molecular evolution of mitochondrial 12S RNA and cytochrome B sequences in the pantherine lineage of Felidae. Mol. Biol. Evol. **12**:690–707.

JELINEK, W. R., and C. W. SCHMID. 1982. Repetitive sequences in eukaryotic DNA and their expression. Annu. Rev. Biochem. **51**:813–844.

JOHNSON, W. E., P. A. DRATCH, J. S. MARTENSON, and S. J. O'BRIEN. 1996. Resolution of recent radiations within three evolutionary lineages of Felidae using mitochondrial restriction fragment length polymorphism variation. J. Mamm. Evol. **3**:97–120.

JOHNSON, W. E., and S. J. O'BRIEN. 1997. Phylogenetic reconstruction of the Felidae using 16SrRNA and Nadh-5 mitochondrial genes. J. Mol. Evol. **44**:S98–S116.

JURKA, J., and P. KLONOWSKI. 1996. Integration of retroposable elements in mammals: selection of target sites. J. Mol. Evol. **43**:685–689.

KIDWELL, M. G., and D. A. LISCH. 1998. Transposons unbound. Nature **393**:22–23.

LAVRENTIEVA, M. V., M. I. RIVKIN, A. G. SHILOV, M. L. KOBETZ, I. B. ROGOZIN, and O. L. SEROV. 1991. B2-like repetitive sequence from the X chromosomes of the American mink (*Mustela vision*). Mamm. Genome **1**:165–170.

MASUDA, R. M., J. V. LOPEZ, J. PECON SLATTERY, N. YUHKI, and S. J. O'BRIEN. 1996. Molecular phylogeny of mitochondrial cytochrome b and 12S rRNA sequences in Felidae: ocelot and domestic cat lineages. Mol. Phylogenet. Evol. **6**:351–365.

MICHEL, D., G. CHATELAIN, C. MAUDUIT, M. BENAHMED, and G. BRUN. 1997. Recent evolutionary acquisition of alternative pre-mRNA splicing and 3′ processing regulations induced by intronic B2 SINE insertion. Nucleic Acid Res. **25**: 3228–3234.

MODI, W. S., and S. J. O'BRIEN. 1988. Quantitative cladistic analysis of chromosomal banding pattern among species in three orders of mammals: homonoid primates, felids and arvicolid rodents. Pp. 215–242 *in* J. P. GUSTAFSON and R. ARPELS, eds. Chromosome structure and function. Plenum Press, New York.

MURPHY, W., S. SHAN, Z.-Q. CHEN, J. PECON SLATTERY, and S. J. O'BRIEN. 1999. Extensive conservation of sex chromosome organization between cat and human revealed by parallel RH mapping. Genome Res. **9**:1223–1230.

NIKAIDO, M., A. ROONEY, and N. OKADA. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest living relatives of whales. Proc. Natl. Acad. Sci. USA **96**:10261–10266.

OHSHIMA, K., and N. OKADA. 1994. Generality of the tRNA origin of short interspersed repetitive elements (SINEs). Characterization of three different tRNA-derived retroposons in the octopus. J. Mol. Biol. **243**:25–37.

OKADA, N. 1991*a*. SINEs: short interspersed repeated elements of the eukaryotic genome. Trends Ecol. Evol. **6**:358–361.

———. 1991*b*. SINEs. Curr. Opin. Genet. Dev. **1**:498–504.

PECON SLATTERY, J., W. E. JOHNSON, D. GOLDMAN, and S. J. O'BRIEN. 1994. Phylogenetic reconstruction of South American felids defined by protein electrophoresis. J. Mol. Evol. **39**:296–305.

PECON SLATTERY, J., and S. J. O'BRIEN. 1998. Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. Genetics **148**:1245–1255.

SCHWARTZ, A., D. C. CHAN, L. G. BROWN, R. ALAGAPPAN, D. PETTAY, C. DISTECHE, B. MCGILLIVRAY, A. DE LA CHAPELLE, and D. PAGE. 1998. Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE-LINE recombination. Hum. Mol. Genet. **7**:1–11.

SHIMAMURA, M., H. YASUE, K. OHSHIMA, H. ABE, H. KATO, T. KISHIRO, M. GOTO, I. MUNECHIKA, and N. OKADA. 1997. Molecular evidence from retroposons that whales form a clade with even-toed ungulates. Nature **388**:666–670.

SMIT, A. F. A. 1996. The origin of interspersed repeats in the human genome. Curr. Opin. Genet. Dev. **6**:743–748.

SMITH, K. D., K. E. YOUNG, C. C. TALBOT JR., and B. J. SCHMECKPEPER. 1987. Repeated DNA of the human Y chromosome. Development **101**(Suppl.):77–92.

TAJIMA, F., and M. NEI. 1984. Estimation of evolutionary distance between nucleotide sequences. Mol. Biol. Evol. **1**: 269–285.

TATOUT, C., L. LAVIE, and J.-M. DERAGON. 1998. Similar target site selection occurs in integration of plant and mammalian retroposons. J. Mol. Evol. **47**:463–470.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGH, and D. G. HIGGINS. 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided quality analysis tools. Nucleic Acids Res. **25**:4876–4882.

WAYNE, R. K., B. VAN VALKENBURGH, and S. J. O'BRIEN. 1991. Molecular distance and divergence time in carnivores and primates. Mol. Biol. Evol. **8**:297–319.

AXEL MEYER, reviewing editor

Accepted January 11, 2000